



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/705, C12N 15/62, C07K 16/28, C12Q 1/68, C12N 5/10		A2	(11) International Publication Number: WO 00/12708 (43) International Publication Date: 9 March 2000 (09.03.00)
(21) International Application Number: PCT/US99/20111 (22) International Filing Date: 1 September 1999 (01.09.99) (30) Priority Data: 60/098,716 1 September 1998 (01.09.98) US 60/098,749 1 September 1998 (01.09.98) US 60/098,750 1 September 1998 (01.09.98) US 60/098,803 2 September 1998 (02.09.98) US 60/098,821 2 September 1998 (02.09.98) US 60/098,843 2 September 1998 (02.09.98) US 60/099,536 9 September 1998 (09.09.98) US 60/099,596 9 September 1998 (09.09.98) US 60/099,598 9 September 1998 (09.09.98) US 60/099,602 9 September 1998 (09.09.98) US (continued after the drawings) (71) Applicant (for all designated States except US): GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): BAKER, Kevin [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). SMITH, Victoria		[AU/US]; 19 Dwight Road, Burlingame, CA 94010 (US). WATANABE, Colin, K. [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). (74) Agents: KRESNAK, Mark, T. et al.; Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: FURTHER PRO POLYPEPTIDES AND SEQUENCES THEREOF			
<pre> CCAATGCCCGCTGCGGTGGTGCAGGGCTCTGGGCTAGTCATGGCGTCCCGCTCTGGAGAC TGCAGACTAAACAGTCATTACTTGTGTTCAAGAGCCTTCTGCTAATCTACACTTTTATTTC TGGATCACTGGCGTTATCCTTCTTGCAGTTGGCATTGGGGCAAGGTGAGCTGGAGAAATTA CTTTCTCTTTTAAATGAGAAGGCCCAATGTCCCTTCGTGCTCATGCTACTGGTACCG TCATTATCTTTTGGGCACTTTGGTGTGTTTGTCTACCTGCCGAGCTTCTGCAATGGATGCTA AACTGTATGCAATGTTTCTGACTCTCGTGTGTTTGGTGGAACTGGTGGCTGCCATCGTAGG ATTGTTTTTCAAGCATGAGATTAAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC AGTATAACTCTACAGGAGATTATAGAAAGCCATGCAGTAGACAAGATCCAAATACGTTGCAT TGTGTGGTGTCAACCAATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAGGATT TCCTAAGAGTTGCTGTAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA ATGAAGGTTGTTTATAAGGTGATGACCAATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA ATTCTCTTGGAGTTGCTTCTTCCAAGTATTGGAATCTTCTCGCTACTGCCNCTCTCG TGCCATAACRAATAACAGTATGAGATAGTGATACCCAATGTATCTGTTGGGCTATTCTCTCT CTACCTTTAAGGACATTTAGGGTCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTG ACAACACTACTTACTGATAGACAAAAAATAACACAGTAGGTTGATTCAATCAAGATGTAT GTAGACCTAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGAGTGGGCTGAT TCAATCAAGATGTATGTTGCTATGTTCTAAGTCCACCTTCTATCCCAATCATGTTAGATCG TTGAAACCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAAATGTAAATGAAGT </pre>			
(57) Abstract			
<p>Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins. The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.</p>			

FURTHER PRO POLYPEPTIDES AND SEQUENCES THEREOF

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides.

5

BACKGROUND OF THE INVENTION

Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.* 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

a qualitative biological activity of a native PRO1311 polypeptide.

In a still further aspect, the invention provides a polypeptide produced by (i) hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO1311 polypeptide having the sequence of amino acid residues from about 45 to about 294, inclusive of Figure 70 (SEQ ID NO:123), or (b) the complement of the DNA molecule of (a), and if the test DNA molecule has at least about an 80% sequence identity, preferably at least about an 85% sequence identity, more preferably at least about a 90% sequence identity, most preferably at least about a 95% sequence identity to (a) or (b), (ii) culturing a host cell comprising the test DNA molecule under conditions suitable for expression of the polypeptide, and (iii) recovering the polypeptide from the cell culture.

10 36. PRO1357

A cDNA clone (DNA64881-1602) has been identified, having homology to nucleic acid encoding the von Ebner minor salivary gland protein that encodes a novel polypeptide, designated in the present application as "PRO1357".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding
15 a PRO1357 polypeptide.

In one aspect, the isolated nucleic acid comprises DNA having at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, most preferably at least about 95% sequence identity to (a) a DNA molecule encoding a PRO1357 polypeptide having the sequence of amino acid residues from about 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID
20 NO:128), or (b) the complement of the DNA molecule of (a).

In another aspect, the invention concerns an isolated nucleic acid molecule encoding a PRO1357 polypeptide comprising DNA hybridizing to the complement of the nucleic acid between about nucleotides 74 or about 137 and about 1525, inclusive, of Figure 71 (SEQ ID NO:127). Preferably, hybridization occurs under stringent hybridization and wash conditions.

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising DNA having at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, most preferably at least about 95% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit No. 203240 (DNA64881-1602) or (b) the complement of the nucleic acid molecule of (a). In a preferred embodiment, the
25 nucleic acid comprises a DNA encoding the same mature polypeptide encoded by the human protein cDNA in
30 ATCC Deposit No. 203240 (DNA64881-1602).

In still a further aspect, the invention concerns an isolated nucleic acid molecule comprising (a) DNA encoding a polypeptide having at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, most preferably at least about 95% sequence
35 identity to the sequence of amino acid residues 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128), or (b) the complement of the DNA of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule having at least 40

nucleotides and produced by hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO1357 polypeptide having the sequence of amino acid residues from 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128), or (b) the complement of the DNA molecule of (a), and, if the DNA molecule has at least about an 80 % sequence identity, preferably at least about an 85 % sequence identity, more preferably at least about a 90 % sequence identity, most preferably at least about a 95 % sequence identity to (a) or (b), isolating the test DNA molecule.

In a specific aspect, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO1357 polypeptide, with or without the N-terminal signal sequence and/or the initiating methionine, or is complementary to such encoding nucleic acid molecule. The signal peptide has been tentatively identified as extending from about amino acid position 1 to about amino acid position 21 in the sequence of Figure 72 (SEQ ID NO:128).

In another aspect, the invention concerns an isolated nucleic acid molecule comprising (a) DNA encoding a polypeptide scoring at least about 80 % positives, preferably at least about 85 % positives, more preferably at least about 90 % positives, most preferably at least about 95 % positives when compared with the amino acid sequence of residues 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128), or (b) the complement of the DNA of (a).

Another embodiment is directed to fragments of a PRO1357 polypeptide coding sequence that may find use as hybridization probes. Such nucleic acid fragments are from about 20 to about 80 nucleotides in length, preferably from about 20 to about 60 nucleotides in length, more preferably from about 20 to about 50 nucleotides in length and most preferably from about 20 to about 40 nucleotides in length and may be derived from the nucleotide sequence shown in Figure 71 (SEQ ID NO:127).

In another embodiment, the invention provides isolated PRO1357 polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a specific aspect, the invention provides isolated native sequence PRO1357 polypeptide, which in certain embodiments, includes an amino acid sequence comprising residues 1 or about 22 to about 484 of Figure 72 (SEQ ID NO:128).

In another aspect, the invention concerns an isolated PRO1357 polypeptide, comprising an amino acid sequence having at least about 80 % sequence identity, preferably at least about 85 % sequence identity, more preferably at least about 90 % sequence identity, most preferably at least about 95 % sequence identity to the sequence of amino acid residues 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128).

In a further aspect, the invention concerns an isolated PRO1357 polypeptide, comprising an amino acid sequence scoring at least about 80 % positives, preferably at least about 85 % positives, more preferably at least about 90 % positives, most preferably at least about 95 % positives when compared with the amino acid sequence of residues 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128).

In yet another aspect, the invention concerns an isolated PRO1357 polypeptide, comprising the sequence of amino acid residues 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128), or a fragment thereof sufficient to provide a binding site for an anti-PRO1357 antibody. Preferably, the PRO1357 fragment retains a qualitative biological activity of a native PRO1357 polypeptide.

In a still further aspect, the invention provides a polypeptide produced by (i) hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO1357 polypeptide having the sequence of amino acid residues from about 1 or about 22 to about 484, inclusive of Figure 72 (SEQ ID NO:128), or (b) the complement of the DNA molecule of (a), and if the test DNA molecule has at least about an 80% sequence identity, preferably at least about an 85% sequence identity, more preferably at least about a 90% sequence identity, most preferably at least about a 95% sequence identity to (a) or (b), (ii) culturing a host cell comprising the test DNA molecule under conditions suitable for expression of the polypeptide, and (iii) recovering the polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO1357 polypeptide. In a particular embodiment, the agonist or antagonist is an anti-PRO1357 antibody.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists of a native PRO1357 polypeptide by contacting the native PRO1357 polypeptide with a candidate molecule and monitoring a biological activity mediated by said polypeptide.

In a still further embodiment, the invention concerns a composition comprising a PRO1357 polypeptide, or an agonist or antagonist as hereinabove defined, in combination with a pharmaceutically acceptable carrier.

37. PRO1244

A cDNA clone (DNA64883-1526) has been identified that encodes a novel polypeptide having homology to Implantation-Associated Protein and designated in the present application as "PRO1244."

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO1244 polypeptide.

In one aspect, the isolated nucleic acid comprises DNA having at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, most preferably at least about 95% sequence identity to (a) a DNA molecule encoding a PRO1244 polypeptide having the sequence of amino acid residues from 1 or about 30 to about 335, inclusive of Figure 74 (SEQ ID NO:130), or (b) the complement of the DNA molecule of (a).

In another aspect, the invention concerns an isolated nucleic acid molecule encoding a PRO1244 polypeptide comprising DNA hybridizing to the complement of the nucleic acid between about residues 96 and about 1013, inclusive, of Figure 73 (SEQ ID NO:129). Preferably, hybridization occurs under stringent hybridization and wash conditions.

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising DNA having at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, most preferably at least about 95% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit No. 203253 (DNA64883-1526), or (b) the complement of the DNA molecule of (a). In a preferred embodiment, the nucleic acid comprises a DNA encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit No. 203253 (DNA64883-1526).

In a still further aspect, the invention concerns an isolated nucleic acid molecule comprising (a) DNA

preferably at least about 90% positives, most preferably at least about 95% positives when compared with the amino acid sequence of residues 31 to about 243, inclusive of Figure 246 (SEQ ID NO:431), or (b) the complement of the DNA of (a).

Another embodiment is directed to fragments of a PRO1550 polypeptide coding sequence that may find use as hybridization probes. Such nucleic acid fragments are from about 20 to about 80 nucleotides in length, preferably from about 20 to about 60 nucleotides in length, more preferably from about 20 to about 50 nucleotides in length, and most preferably from about 20 to about 40 nucleotides in length.

In another embodiment, the invention provides isolated PRO1550 polypeptide encoded by any of the isolated nucleic acid sequences hereinabove defined.

In a specific aspect, the invention provides isolated native sequence PRO1550 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 31 to 243 of Figure 246 (SEQ ID NO:431).

In another aspect, the invention concerns an isolated PRO1550 polypeptide, comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, most preferably at least about 95% sequence identity to the sequence of amino acid residues 31 to about 243, inclusive of Figure 246 (SEQ ID NO:431).

In a further aspect, the invention concerns an isolated PRO1550 polypeptide, comprising an amino acid sequence scoring at least about 80% positives, preferably at least about 85% positives, more preferably at least about 90% positives, most preferably at least about 95% positives when compared with the amino acid sequence of residues 31 to 243 of Figure 246 (SEQ ID NO:431).

In yet another aspect, the invention concerns an isolated PRO1550 polypeptide, comprising the sequence of amino acid residues 31 to about 243, inclusive of Figure 246 (SEQ ID NO:431), or a fragment thereof sufficient to provide a binding site for an anti-PRO1550 antibody. Preferably, the PRO1550 fragment retains a qualitative biological activity of a native PRO1550 polypeptide.

In a still further aspect, the invention provides a polypeptide produced by (i) hybridizing a test DNA molecule under stringent conditions with (a) a DNA molecule encoding a PRO1550 polypeptide having the sequence of amino acid residues from about 31 to about 243, inclusive of Figure 246 (SEQ ID NO:431), or (b) the complement of the DNA molecule of (a), and if the test DNA molecule has at least about an 80% sequence identity, preferably at least about an 85% sequence identity, more preferably at least about a 90% sequence identity, most preferably at least about a 95% sequence identity to (a) or (b), (ii) culturing a host cell comprising the test DNA molecule under conditions suitable for expression of the polypeptide, and (iii) recovering the polypeptide from the cell culture.

124. Additional Embodiments

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

5 In another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences, wherein those probes may be derived from any of the above or below described nucleotide sequences.

10 In other embodiments, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99%

sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein or the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein, or (b) the complement of the DNA molecule of (a).

5 In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89%
10 sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more
15 preferably at least about 99% sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s)
20 of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence that may find use as, for example, hybridization probes or for encoding fragments of a PRO polypeptide that may optionally
25 encode a polypeptide comprising a binding site for an anti-PRO antibody. Such nucleic acid fragments are usually at least about 20 nucleotides in length, preferably at least about 30 nucleotides in length, more preferably at least about 40 nucleotides in length, yet more preferably at least about 50 nucleotides in length, yet more preferably at least about 60 nucleotides in length, yet more preferably at least about 70 nucleotides in length, yet more preferably at least about 80 nucleotides in length, yet more preferably at least about 90 nucleotides in
30 length, yet more preferably at least about 100 nucleotides in length, yet more preferably at least about 110 nucleotides in length, yet more preferably at least about 120 nucleotides in length, yet more preferably at least about 130 nucleotides in length, yet more preferably at least about 140 nucleotides in length, yet more preferably at least about 150 nucleotides in length, yet more preferably at least about 160 nucleotides in length, yet more preferably at least about 170 nucleotides in length, yet more preferably at least about 180 nucleotides in length,
35 yet more preferably at least about 190 nucleotides in length, yet more preferably at least about 200 nucleotides in length, yet more preferably at least about 250 nucleotides in length, yet more preferably at least about 300 nucleotides in length, yet more preferably at least about 350 nucleotides in length, yet more preferably at least

about 400 nucleotides in length, yet more preferably at least about 450 nucleotides in length, yet more preferably at least about 500 nucleotides in length, yet more preferably at least about 600 nucleotides in length, yet more preferably at least about 700 nucleotides in length, yet more preferably at least about 800 nucleotides in length, yet more preferably at least about 900 nucleotides in length and yet more preferably at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably

at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence scoring at least about 80% positives, preferably at least about 81% positives, more preferably at least about 82% positives, yet more preferably at least about 83% positives, yet more preferably at least about 84% positives, yet more preferably at least about 85% positives, yet more preferably at least about 86% positives, yet more preferably at least about 87% positives, yet more preferably at least about 88% positives, yet more preferably at least about 89% positives, yet more preferably at least about 90% positives, yet more preferably at least about 91% positives, yet more preferably at least about 92% positives, yet more preferably at least about 93% positives, yet more preferably at least about 94% positives, yet more preferably at least about 95% positives, yet more preferably at least about 96% positives, yet more preferably at least about 97% positives, yet more preferably at least about 98% positives and yet more preferably at least about 99% positives when compared with the amino acid sequence of a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist

or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Figure 1 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO1560 (UNQ767) cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA19902-1669". The start and stop codons are shown in bold and underlined font.

 Figure 2 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 1.

10 Figure 3 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO444 (UNQ328) cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA26846-1397". The start and stop codons are shown in bold and underlined font.

 Figure 4 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 3.

15 Figure 5 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO1018 (UNQ501) cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA56107-1415". The start and stop codons are shown in bold and underlined font.

 Figure 6 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 5.

20 Figure 7 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO1773 (UNQ835) cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA56406-1704". The start and stop codons are shown in bold and underlined font.

 Figure 8 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in Figure 7.

25 Figure 9 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO1477 (UNQ747) cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA56529-1647". The start and stop codons are shown in bold and underlined font.

 Figure 10 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 9.

30 Figure 11 shows a nucleotide sequence (SEQ ID NO:16) of a native sequence PRO1478 (UNQ748) cDNA, wherein SEQ ID NO:16 is a clone designated herein as "DNA56531-1648". The start and stop codons are shown in bold and underlined font.

 Figure 12 shows the amino acid sequence (SEQ ID NO:17) derived from the coding sequence of SEQ ID NO:16 shown in Figure 11.

35 Figure 13 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO831 (UNQ471) cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA56862-1343". The start and stop codons are shown in bold and underlined font.

Figure 59 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO1249 (UNQ632) cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA62809-1531". The start and stop codons are shown in bold and underlined font.

Figure 60 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:100 shown in Figure 59.

- 5 Figure 61 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO1315 (UNQ681) cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA62815-1578". The start and stop codons are shown in bold and underlined font.

Figure 62 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 61.

- 10 Figure 63 shows a nucleotide sequence (SEQ ID NO:110) of a native sequence PRO1549 (UNQ782) cDNA, wherein SEQ ID NO:110 is a clone designated herein as "DNA62845-1684". The start and stop codons are shown in bold and underlined font.

Figure 64 shows the amino acid sequence (SEQ ID NO:111) derived from the coding sequence of SEQ ID NO:110 shown in Figure 63.

- 15 Figure 65 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO1430 (UNQ736) cDNA, wherein SEQ ID NO:115 is a clone designated herein as "DNA64842-1632". The start and stop codons are shown in bold and underlined font.

Figure 66 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in Figure 65.

- 20 Figure 67 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO1374 (UNQ711) cDNA, wherein SEQ ID NO:117 is a clone designated herein as "DNA64849-1604". The start and stop codons are shown in bold and underlined font.

Figure 68 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in Figure 67.

- 25 Figure 69 shows a nucleotide sequence (SEQ ID NO:122) of a native sequence PRO1311 (UNQ677) cDNA, wherein SEQ ID NO:122 is a clone designated herein as "DNA64863-1573". The start and stop codons are shown in bold and underlined font.

Figure 70 shows the amino acid sequence (SEQ ID NO:123) derived from the coding sequence of SEQ ID NO:122 shown in Figure 69.

- 30 Figure 71 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO1357 (UNQ706) cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA64881-1602". The start and stop codons are shown in bold and underlined font.

Figure 72 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 71.

- 35 Figure 73 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO1244 (UNQ628) cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA64883-1526". The start and stop codons are shown in bold and underlined font.

5'-CCTGTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCC-3' (SEQ ID NO:126)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1311 gene using the probe oligonucleotide and one of the PCR primers.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 195-197, and a stop signal at nucleotide positions 1077-1079 (Figure 69; SEQ ID NO:122). The predicted polypeptide precursor is 294 amino acids long has a calculated molecular weight of approximately 33,211 daltons and an estimated pI of approximately 5.35. Additional features include: a signal sequence at about amino acids 1-44; possible transmembrane domains at about amino acids 22-42, 57-85, 94-116, and 230-257; potential N-glycosylation sites at about amino acids 118-121, 1899-192, and 230-233; potential tyrosine kinase phosphorylation sites at about amino acids 3-11 and 129-136; potential N-myristoylation sites at about amino acids 80-85, 109-114, 180-185, 218-223, 248-253, 276-281, 285-290, and 287-292; and a cell attachment sequence at about amino acids 3-5.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 70 (SEQ ID NO:123), evidenced some homology between the PRO1311 amino acid sequence and the following Dayhoff sequences: AF065389_1, AF053455_1, CD63_HUMAN, A15_HUMAN, AF043906_1, C151_HUMAN, AF053453_1, AF054838_1, P_R91446, and CD82_HUMAN.

Clone DNA64863-1573 was deposited with the ATCC on September 9, 1998, and is assigned ATCC deposit no. 203251.

EXAMPLE 39: Isolation of cDNA clones Encoding Human PRO1357

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated Incyte EST cluster sequence no. 69537. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56034.

In light of the sequence homology between the DNA56034 sequence and an EST sequence contained within the Incyte EST clone no. 936239, the Incyte EST clone no. 936239 was purchased and the cDNA insert was obtained and sequenced. The sequence of this cDNA insert is shown in Figure 71 and is herein designated as DNA64881-1602.

Clone DNA64881-1602 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 74-76 and ending at the stop codon at nucleotide positions 1526-1528 (Figure 71). The predicted polypeptide precursor is 484 amino acids long (Figure 72). The full-length PRO1357 protein

shown in Figure 72 has an estimated molecular weight of about 52,468 daltons and a pI of about 7.14. Analysis of the full-length PRO1357 sequence shown in Figure 72 (SEQ ID NO:128) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 21, potential N-glycosylation sites from about amino acid 48 to about amino acid 51, from about amino acid 264 to about amino acid 267 and from about amino acid 401 to about amino acid 404, a glycosaminoglycan attachment site from about amino acid 412 to about amino acid 415 and an amino acid sequence block having homology to the LBP/BPI/CETP family of proteins from about amino acid 407 to about amino acid 457. Clone DNA64881-1602 has been deposited with ATCC on September 9, 1998 and is assigned ATCC deposit no. 203240.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 72 (SEQ ID NO:128), evidenced significant homology between the PRO1357 amino acid sequence and the following Dayhoff sequences: MMU46068_1, S17447, MMU1_1, BPI_RABIT, P_W16808, P_R21844, PSP_MOUSE, HSLBPEX1_1 and BTU79413_1.

EXAMPLE 40: Isolation of cDNA clones Encoding Human PRO1244

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, designated cluster no. 7874. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA databases (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA; Genentech, South San Francisco, CA) to identify existing homologies. One or more of the ESTs was derived from a library constructed from tissue of the corpus cavernosum. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated "DNA56011".

In light of the sequence homology between the DNA56011 sequence and an EST sequence contained within Incyte EST No. 3202349, the EST clone no. 3202349 was purchased and the cDNA insert was obtained and sequenced. The sequence of this cDNA insert is shown in Figure 73 (SEQ ID NO:129) and is herein designated "DNA64883-1526".

The full length clone shown in Figure 73 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 9-11 and ending at the stop codon found at nucleotide positions 1014-1016 (Figure 73; SEQ ID NO:129). The predicted polypeptide precursor (Figure 74, SEQ ID NO:130) is 335 amino acids long. PRO1244 has a calculated molecular weight of approximately 38,037 daltons and an estimated pI of approximately 9.87. Other features include a signal peptide at about amino acids 1-29; transmembrane domains at about amino acids 183-205, 217-237, 271-287, and 301-321; potential N-glycosylation sites at about amino acids 71-74, and 215-218; and a cell attachment sequence at about amino acids 150-152.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence

FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGA
CACCTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACC
TTGATCCAAGCCACCCTCAGTCCCCTGACAGTTCTCATCCTCGGCCCCAAAAGTCATCAAAGA
AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGC
TCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC
GTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAA
GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA
ACACGCCCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC
CGCATGGACACCAGTGCAAGTGGCCCCACCCGCTGGTCCTCAGTGACTGTGCCACCAGCCA
TGGGAGCCTGCGCATCCAAGTGTGTATAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGC
AGGTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG
ATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTT
CCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATT
AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAAC
TCTGCAGCTTCCCTGACAATGCCACCCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCA
GGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCCTGTTGG
ACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG
GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT
TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAAGTGTTTCCCTCCA
GTGAAGCCCTCCGCCCTTTGTTTACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTAC
ACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT
GAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCC
ACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTATTGGTG
AAGGCCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCC
AGCCTCCTTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAG
GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT
CAATAAACACTTGCCTGTGAAAAA

FIGURE 72

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881
><subunit 1 of 1, 484 aa, 1 stop
><MW: 52468, pI: 7.14, NX(S/T): 3
MAGPWTFLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAM
REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL
VKTIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN
LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYL
GAKLLDSQGVTKWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVL
PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEAL
RPLFTLGLIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSIL
LPNQNGKLRSVGPVSLVKALGFEEAESSLTkdALVLTpASLWKpSSPVSQ
```

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457